

EP 30124 (2)

PCTWORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

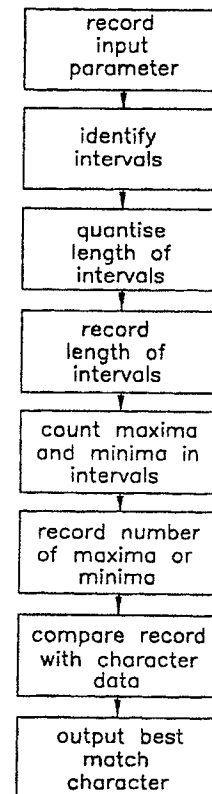
INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁷ : G06K 9/22, 9/00		A1	(11) International Publication Number: WO 00/62244
			(43) International Publication Date: 19 October 2000 (19.10.00)
(21) International Application Number: PCT/GB00/01340 (22) International Filing Date: 14 April 2000 (14.04.00) (30) Priority Data: 9908462.6 14 April 1999 (14.04.99) GB (71) Applicant (for all designated States except US): NEW TRANS- DUCERS LIMITED [GB/GB]; 37 Ixworth Place, London SW3 3QH (GB). (72) Inventors; and (75) Inventors/Applicants (for US only): WALKER, Ross [GB/GB]; Station House, Station Road, Cliddesden, Hamp- shire RG25 2NL (GB). KAVANAGH, Sean [GB/GB]; 65A High Street, Huntingdon PE29 3DN (GB). FERIKIDIS, Charalampos [DE/DE]; Im Wieh 7, D-30890 Barsinghausen OT Barrigsen (DE). (74) Agent: MAGUIRE BOSS; 5 Crown Street, St. Ives, Cam- bridgeshire PE17 4EB (GB).		(81) Designated States: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG). Published With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.	

(54) Title: HANDWRITING CODING AND RECOGNITION

(57) Abstract

A coding and recognition method for handwriting identifies intervals between occurrences of an input parameter, such as pen position and velocities, crossing a given value, and recording the lengths of the intervals. The number of maxima and minima in the intervals are then recorded.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

5 TITLE: HANDWRITING CODING AND RECOGNITION10 DESCRIPTION

15 This invention relates to handwriting coding and recognition, and in particular for a system that uses a coding method to generate data for recognition.

 The problem of recognising handwriting has attracted much attention in recent years, particularly for use as a
20 system for computer input, especially for small notebook-type computer systems and so called "palm-top" computer systems.

 Handwriting recognition can be divided up into distinct types of recognition problem. The main
25 distinction is between continuous writing, where there are no gaps between letters and script where the characters are discrete. Since there are always gaps between words in continuous writing it is logical for a recognition

system to recognise whole words which are easier to isolate and probably easier to recognise.

Another important distinction in a hand writing recognition scheme is whether the letters are produced as they are being analysed or if the characters have been produced beforehand. In the former case more information will be available.

In a different field, it is known to time encode speech using a Time Encoded Signal Processing and Recognition (TESPAR) coder. Aspects of this coder are described in GB 2020517, GB 2084433, GB 2162024, GB 2162025, GB 2187586, GB 2179 183, WO 92/15089, WO97/31368, WO97/45831 and WO98/08188.

According to the first aspect of the invention there is provided a coding method for handwriting, comprising the steps of

recording the variation of an input parameter,

identifying the intervals between the occurrences of the input parameter crossing a given value, and quantising the lengths of the intervals,

identifying the number of complex zeroes of the input parameter, up to a predetermined rank, in the said intervals, and

recording the quantised lengths of the intervals and a measure of the said number complex zeroes up to a predetermined rank as a representation of the variation of the input parameter.

A predetermined rank of 1 has been found to give good

results. In this case the method records the number of first rank zeroes, i.e. positive minima or negative maxima. This information may provide sufficient detail for useful characterisation without requiring excessive calculation.

The method thus parameterises the shape of the input parameter function. If the parameter rises smoothly to a maximum and then falls smoothly to the next zero, there will be no positive minimal so said number will be zero.

10 If the function has an "M" shape, rising to a maximum, falling to a minimum and then rising to another maximum before passing through zero, then there will be one positive minimum so the said number will be one.

Thus, the number parameterises the number of 15 oscillations of the input parameter between zeroes.

The reason that the positive minima or negative maxima are known as complex zeroes of a function is that they correspond to zeroes of the function for complex number inputs to the function. The first rank zeroes 20 occur at real values being the real values of the complex numbers for which the function has a value zero.

The coding method may be a TESPAP method.

The method may further comprise the step of generating a code number taking one of a set of 25 predetermined values representing the duration of the interval and number of maxima and minima for at least some of the said intervals.

The code numbers may be further parameterised. In

one approach an S matrix may be calculated. The S matrix records the number of instances of each code number in the recorded variation of the input parameter. Alternatively or additionally an A matrix recording the number of 5 instances of a first code number following a second code number with a predetermined lag may be calculated. A further alternative is to calculate a DZ matrix recording the number of instances of amplitude, length of interval and number of maxima and minima increasing, decreasing or 10 staying the same in the next epoch.

The S, A and/or DZ matrices may be stored or evaluated.

The above approach may be used to record at least two input parameters. The input parameters may be selected 15 from x and y coordinates, speed, one or more coordinates of velocity (e.g. speed and angle, or x coordinate and y coordinate), angular velocity, or radius of curvature.

The lengths of the intervals may be measured in time, i.e. the recorded length represents the time between 20 successive crossings of the given value. However, other parameters, such as the x coordinate, could also be used. In the latter case the recorded length is a distance. TESPAR works on single-valued function of a variable.

The said given value is preferably a predetermined 25 value. The value may be a zero value, particularly for a coordinate such as the vertical (y) component of velocity which passes through zero. The given value may also be a value corresponding to a median value of the handwriting.

This approach can be used to parameterise the y-coordinate position, for example, in which case the value can be the centre line of the handwriting, i.e. the median y-coordinate assuming that the handwriting is written in the x direction.

Preferably, a writing tablet is provided for capturing the input parameter or parameters. The writing tablet may also output position data which can be converted into other parameters using a computer.

10 Alternatively, the method may be used on handwriting that is already written, by scanning the handwriting, and then generating the input parameters. This may be done by using known techniques for vectorising bitmap data.

The first aspect of the invention also envisages a method of handwriting recognition, including the steps of coding the handwriting as described above, comparing the coded representations with representations corresponding to a number of different characters to find the best match and outputting the character corresponding to the best match.

The comparing step may be carried out using a matrix method in which the correlations between matrices representing the coded data and matrices representing each of the said plurality of different characters are compared. Alternatively, a neural network approach to the comparison could be used. Either approach may require a training period to generate suitable matrices.

According to a second aspect of the invention there

is provided a handwriting coding apparatus, comprising

a data input means for measuring the variation of an input parameter,

an identifying means for identifying the intervals between the occurrences of the input parameter crossing a given value, and quantising the intervals, identifying the number of maxima and minima of the input parameter in said intervals, and

a recording means for recording the quantised intervals and the said number of maxima and minima as a representation of the variation of the input parameter.

Preferably, the above approach is used to record at least two input parameters. The input parameters may be selected from x and y coordinates, speed, one or more coordinates of velocity (e.g. speed and angle, or x coordinate and y coordinate), angular velocity, or radius of curvature.

Preferably, a writing tablet is provided for capturing the input parameters. The writing tablet may be of a conventional form. Alternatively, the method may be used on handwriting that is already written, by providing a scanner for scanning the handwriting, and then generating the input parameters in the data input means by using known techniques, such as those used for vectorising bitmap data.

The second aspect of the invention may also provide a handwriting recognition apparatus, including a handwriting coding apparatus as described above, wherein the apparatus

further comprises a processing means adapted to compare the coded representations of the characters with representations corresponding to a number of different characters to find the best match and outputting the 5 character corresponding to the best match.

Preferably, the identifying means and the processing means are both the central processing unit of a computer.

According to a third aspect of the invention there is provided a computer program recorded on a data carrier, 10 operable to control a computer having a data store, an input device and a processor, the computer program operable to cause the said computer to carry out in cooperation with the input device a method for coding handwriting comprising the steps of

15 recording the variation of an input parameter, identifying the intervals between the occurrences of the input parameter crossing a given value, and quantising the intervals,

identifying the number of first rank zeroes of the 20 input parameter in said intervals, and

recording the quantised intervals and the said number of first rank zeroes as a representation of the variation of the input parameter.

The input device may be a writing tablet or a 25 scanner, for example.

The computer program recorded on a data carrier according to the third aspect preferably also causes the computer to carry out the step of comparing the coded

representation with representations corresponding to a number of different characters to find the best match and outputting the character corresponding to the best match.

For a better understanding of the invention an embodiment will now be described, purely by way of example, with reference to the accompanying drawings, in which

Figure 1 shows a schematic diagram of an apparatus according to the invention,

10 Figure 2 shows a flow diagram of a coding method in accordance with the invention,

Figure 3 shows a handwriting recognition scheme in accordance with the invention for recognising static text, and

15 Figure 4 shows a handwriting recognition scheme in accordance with the invention for recognising dynamic text.

Referring to Fig.1, a writing tablet (1) records the x- and y- coordinates of a stylus (2) used to write thereupon. The tablet is connected to a computer (3) 20 having a display (5), a processor (9) and a data store (11). The computer is connected to the tablet by a connector (13).

A program is stored in the data store (11) to cause 25 the system to carry out a data coding method, schematically shown in Fig.2.

Firstly, data corresponding to input coordinates is captured by the tablet (1) which records the position of

the stylus as a user writes with the stylus (2) on the tablet (1). The tablet (1) outputs the x- and y-coordinates of the stylus (2). The x and y components of the velocity are then calculated by the computer. The
5 method then continues using the y- coordinate of position and the x- and y- coordinates of velocity as the three input parameters. These input parameters are not exclusive, and for some arrangements other sets of input parameters may be more suitable.

10 The computer then calculates the centre line of the writing, and uses that as the special value for the y-coordinate input parameter. The zero value of the x and y components of the velocity is used for those components. For convenience, all these given values will be referred
15 to as "zero values" hereinafter.

It is also possible to eliminate any offsets by differentiating the input once, then parameterising the input.

The data is parameterised by quantising the time
20 interval between zero crossings, i.e. the time between the consecutive occasions that the input data parameter crosses the zero value. Then, the number of first rank zeroes in each time interval is recorded. The number of first rank zeroes is the number of negative maxima or
25 positive minima. The interval and number of zeroes is then coded using a single integer between 1 and 28. The integer 1 for example corresponds to an interval of length up to a predetermined time and no first rank zeroes. Each

range of intervals and number of zeroes is assigned to a predetermined integer, as set out in GB 2162024.

The complete writing sample is then parameterised using an "S" matrix, as used in the TESPAP method. The "S" matrix records the number of instances of each code number in the data sample.

The method compares the recorded values with values for each possible character of the alphabet and numerals 1 to 9, to record which character has been written. The character that gives the best match for each of the three input parameters is output, and the computer then goes on to evaluate the next character.

In the specific example the method includes the so-called "TESPAR" method. A speech encoder known as TESPAP is known from GB2162024. The method of the invention may carry out the coding using the TESPAP method and specific hardware adapted to implement the method. A brief summary of TESPAP will now be given.

Any band limited signal can be represented exactly in 20 terms of its real and complex zeros, such that

$$f(t) = \prod_i (t - \tau_i) \quad (1)$$

where τ_i are the zeros of the function. Real zeros correspond to times where the function $f(t)$ crosses the zero line.

Complex zeros occur in conjugate pairs for a real function and can be classified by considering the effect

of differentiation with respect to time. The rank of a complex zero is defined as the number of differentiations required to produce a real zero at the same time ordinate as the complex zero. Thus positive minima (minima at 5 positive values) and negative maxima (maxima at negative values) of the waveform correspond to first rank complex zeros because these stationary points become zero crossings when the waveform is differentiated once.

The waveform in TESPAP may be divided up into 10 sections using the real zeros as the boundary points between sections. Each interval between two real zeros is then referred to as an epoch. This is not the only method available, but it is simple and usually effective.

Once the epochs have been defined some relevant 15 information needs to be recorded about them. This conventionally includes the length of the interval (duration) of the epoch; its maximum amplitude and the number of first rank zeros. At this point the duration is usually represented as the nearest integer number of 20 samples of the waveform rather than the exact time duration. The result is a matrix of values of size (3 x number of epochs) containing some of the information contained in the waveform.

It should be noted that at this point an 25 approximation to the original waveform could still be regenerated. It would not be exact since the duration of each epoch has been quantised and only the presence of first rank zeros has been noted. The position of the zeros

and the existence of higher rank zeros are not recorded. However, there is sufficient information stored to allow a good approximation to the original waveform to be generated in the sense that a regenerated speech signal is easily understood.

Once the epochs have been defined and their parameters recorded the time encoding can be done. There are many schemes for encoding a signal once it has been stored in epoch form as described above (called the natural TESPAR stream). The general idea however is to take each epoch in turn, or several epochs at a time, and produce a code number depending upon the information stored in the epoch(s). A common encoding scheme considers the duration and number of first rank zeroes and returns a single code number, dependent upon these values, ranging between 1 and 28.

When the epochs have been encoded some more information has been lost, but a signal can still be regenerated. However, at this point the aim is not usually to regenerate the signal but to produce a compact format that will allow the waveform to be characterised and compared with other waveforms. To this end the code numbers for each epoch are combined into one of a variety of possible matrices. The more common types of matrix are described below.

The matrix is actually a vector of length equal to the number of TESPAR codes used to describe all the different types of epoch. Each element of the vector

contains the number of times that an epoch with that code occurred in the signal. The resulting vector contains information about the content of the signal but no information about the ordering of the epochs, which means
5 the signal cannot be regenerated without extra information.

The 'S' matrix can be further refined. Rather than incrementing an element of the vector by one for each epoch with the appropriate code the element can be
10 incremented by an amount that depends upon some characteristic of the epoch. This leads to the duration weighted 'S' matrix, where the elements are incremented by an amount dependent on the duration of the epoch and the amplitude weighted 'S' matrix where the maximum epoch
15 amplitude is used as weighting. In addition the final matrix itself can be modified. For example the number in each element of the 'S' matrix can be doubled to produce a matrix that is more heavily weighted by the number of epochs that occur in the waveform.

20 The 'A' matrix is a two dimensional matrix that is generated by considering pairs of epochs in turn. These can be adjacent epochs or they can have a specified separation (called the lag). Each epoch in the pair will have a TESPAC code associated with it which gives two
25 ordinates that specify the element of the matrix that is to be incremented. The 'A' matrix is similar to the 'S' matrix but some information as to the ordering of the epochs is retained. An example of this matrix is shown in

figure 2. As with the 'S' matrix the elements of the 'A' matrix can be weighted according to some combination of epoch parameters.

It is possible to extend the idea of the 'A' matrix 5 by considering groups of three epochs at a time to produce a three dimensional matrix. Indeed this can be extended up to the number of epochs in the signal at which point there is one non-zero element in a matrix with the same number of dimensions as there are epochs in the waveform.

10 The 'DZ' matrix is generated by considering pairs of epochs and looking at how they change. Specifically the maximum amplitude, duration and number of complex first rank zeros are examined in each epoch. Each of these parameters may increase, decrease or stay the same which 15 gives 27 possible combinations and hence a 27 element vector. Thus the 'DZ' matrix looks similar to an 'S' matrix. Since duration and amplitude are nearly continuous it is usual to specify a range of changes for these values that are taken to be the same for the purposes of encoding 20 the 'DZ' matrix.

To illustrate how an end-point finding approach might work consider a speech waveform with individual words separated by silence. The signal is divided up into short time segments (say 20ms) and each time segment encoded 25 into an 'S' matrix. Each element of these 'S' matrices is then summed up to give one number per time slice reflecting how many epochs are present in each time slice. Note that the silence may include a lot of small amplitude

epochs as unavoidable background noise. In this case some form of suppression is required such as ignoring all epochs with maximum amplitude less than a certain value. This simple graph will then have low values for regions of 5 silence and high values in regions where a word is being spoken, which allows the end points of the words to be estimated.

Once the appropriate section of a waveform has been located and encoded in some form of matrix it usually 10 needs to be categorised as an example of a previously known signal. The two most common approaches are to use archetypal matrices or to use a neural net.

Given a set of examples of a particular signal an archetypal matrix can be generated by simply adding the 15 matrices together and dividing by the number of matrices used. To test to see if a signal is the same as the one represented by the archetype the signal needs to be encoded in the same matrix format as the archetype and then compared, usually by finding the correlation score. 20 In this way many archetypes can be included and ranked as to which are most likely to be the same as the signal.

If better discrimination is required than that afforded by using archetype matrices a neural net may be used. Since a signal of any length will always produce a 25 matrix of the same size when it is time encoded (assuming the same matrix is used, of course) a neural net can be designed with a fixed number of inputs corresponding to the number of elements in the matrix. With sufficient

training examples the neural net can be trained to recognise a number of standard signals.

If the handwriting is analysed as it is being written (or the relevant data is recorded) the position of the pen as a function of time is available. The (x,y) co-ordinates provide two waveforms that vary with time and are orthogonal. These can be encoded separately to provide two separate matrices or they can be used to derive other parameters. An example would be differentiating $x(t)$ and $y(t)$ once to produce the velocities. In each case the waveform chosen has to have a "zero" value that the waveform crosses so that the epochs can be defined. In the case of the velocities this could be the actual zero value whereas for the y ordinate a line through the middle of the text would be required.

An additional complication arises when the pen is lifted off the surface of whatever is being written on. This could happen for example when the crossbar on a "t" is written after the vertical line has been drawn. Such an event will be characterised by gaps in the plots at certain times. Possible methods of incorporating these gaps include setting the value of the functions to zero over the period of the gaps or simply joining the separate segments together. In either case there will be discontinuous changes in the waveforms. However, this should not be a problem for time encoding.

Figure 3 illustrates one possible implementation of a system to recognise text as it is being written. In the

first step 31, data of the position and velocity of the pen on a writing tablet is obtained. The y coordinate of the writing implement about a mean level and the velocity of the implement in the x direction as a function of time 5 are then extracted from the data in the next step 33. These two waveforms are then time encoded 35 using the TESPAP encoding scheme and compared with a store of previous examples. The closest match between the known examples and the input data is then output.

10 A specific application of the above recognition example is signature recognition. The person in question signs their name on a tablet that records the position as a function of time. This signal can then be time-encoded and compared to known time-encoded signatures to verify 15 the identity of that person.

The recognition method may also be applied to static text. When the text to be recognised has already been written the information as to how it was written is not available. Hence this case provides a subset of the 20 information available in the dynamic, case where the text is analysed as it is written.

One representation of written script is as a bitmap in which the pixels are either black (foreground colour) or white (background colour). This gives a two-dimensional 25 array of points each of which takes the value 0 or 1. This can be thought of as an infinitely clipped two dimensional waveform. (An alternative would be to assign a number to each colour that can be represented to produce a function

that varies over a range of values.) When this waveform has been generated it can be encoded. One method is to encode each column (or row) into matrix form and then use the resulting set of matrices as the input for a neural 5 net. Alternatively the array could be turned into a one-dimensional vector by writing each row (or column) sequentially. The resulting vector could then be encoded to produce a single matrix that could be used to characterise the image. This example describes one method 10 of implementing a system to recognise scanned images of static text. The flow diagram for this system is shown in Figure 4. Initially, an image of the text to be classified is scanned (41) to render it in electronic form and stored as a two dimensional bitmap (43). This image is then read 15 (45) both one row after another and one column after another to produce two long waveforms corresponding to the row and column scans. A pixel that is clear is represented by minus 1 and a filled pixel by plus 1. This produces a rectangular waveform with well defined zero crossing 20 points.

The resulting encoded waveforms are then compared (47) with previously determined examples to assign probabilities to each. The most likely is then returned (49) as the recognised text.

CLAIMS

1. A coding method for handwriting, comprising the steps of
recording the variation of an input parameter,
5 identifying the intervals between the occurrences of the input parameter crossing a given value, and quantising the lengths of the intervals,
identifying the number of complex zeroes, of the input parameter up to a predetermined rank, in the said
10 intervals, and
recording the quantised lengths of the intervals and a measure of the said number of complex zeroes up to a predetermined rank as a representation of the variation of the input parameter.
- 15 2. A method according to claim 1 wherein the predetermined rank is 1 and the steps of recording a representation of the variation of the input parameter records the number of first rank zeroes.
3. A method according to claim 2 wherein the coding
20 method is a TESPAS method.
4. A method according to any preceding claim further comprising the step of generating a code number taking one of a set of predetermined values representing the duration of the interval and the said number of complex zeroes for
25 at least some of the said intervals.
5. A method according to claim 4 further comprising the step of calculating an S matrix recording the number of instances of each code number in the recorded variation of

the input parameter.

6. A method according to claim 4 or 5 further comprising the step of calculating an A matrix recording the number of instances of a first code number following a second 5 code number with a predetermined lag.

7. A method according to any of claims 4 to 6 further comprising the step of calculating a DZ matrix recording the number of instances of amplitude, length of interval and number of maxima and minima increasing, decreasing or 10 staying the same in the next epoch.

8. A method according to any preceding claim wherein at least two input parameters are recorded.

9. A method according to any preceding claim wherein the intervals are time intervals so that the recorded lengths 15 represent the time between successive crossings of the given value.

10. A method according to any preceding claim wherein the given value is a median value of the handwriting.

11. A method according to any preceding claim including 20 capturing the input parameter or parameters with a writing tablet.

12. A method according to any of claims 1 to 10 including scanning the handwriting, and then generating the input parameters.

25 13. A method of handwriting recognition, including the steps of

coding the handwriting according to any preceding claim,

comparing the coded representations with representations corresponding to a number of different characters to find the best match and

outputting the character corresponding to the best match.

14. A handwriting coding apparatus, comprising

a data input means for measuring the variation of an input parameter,

an identifying means for identifying the intervals between the occurrences of the input parameter crossing a given value, and quantising the intervals, and identifying the number of complex zeroes up to a predetermined rank of the input parameter in said intervals, and

a recording means for recording the quantised intervals and a measure of the said number of complex zeroes as a representation of the variation of the input parameter.

15. A handwriting coding apparatus according to claim 14 wherein the predetermined rank is one.

16. A handwriting coding apparatus according to claim 14 or 15 further comprising a writing tablet for capturing the input parameters.

17. A handwriting recognition apparatus, including a handwriting coding apparatus according to any of claims 14 or 16 and further comprising

a processing means adapted to compare the coded representations of the characters with representations corresponding to a number of different characters to find

the best match and outputting the character corresponding to the best match.

18. A handwriting recognition apparatus according to claim 17 wherein the identifying means and the processing means are both the central processing unit of a computer.

19. A computer program for controlling a computer having a data storage means and a processor, connected to an input means for providing handwriting data input, the computer program operable to cause the said computer to carry out in cooperation with the input means a method for coding handwriting comprising the steps of

recording the variation of an input parameter,

identifying the intervals between the occurrences of the input parameter crossing a given value, and quantising the intervals,

identifying the number of complex zeroes of first rank of the input parameter in said intervals, and

recording the quantised intervals and a measure of the said number of coupler zeroes as a representation of the variation of the input parameter.

Fig 1.

1/3

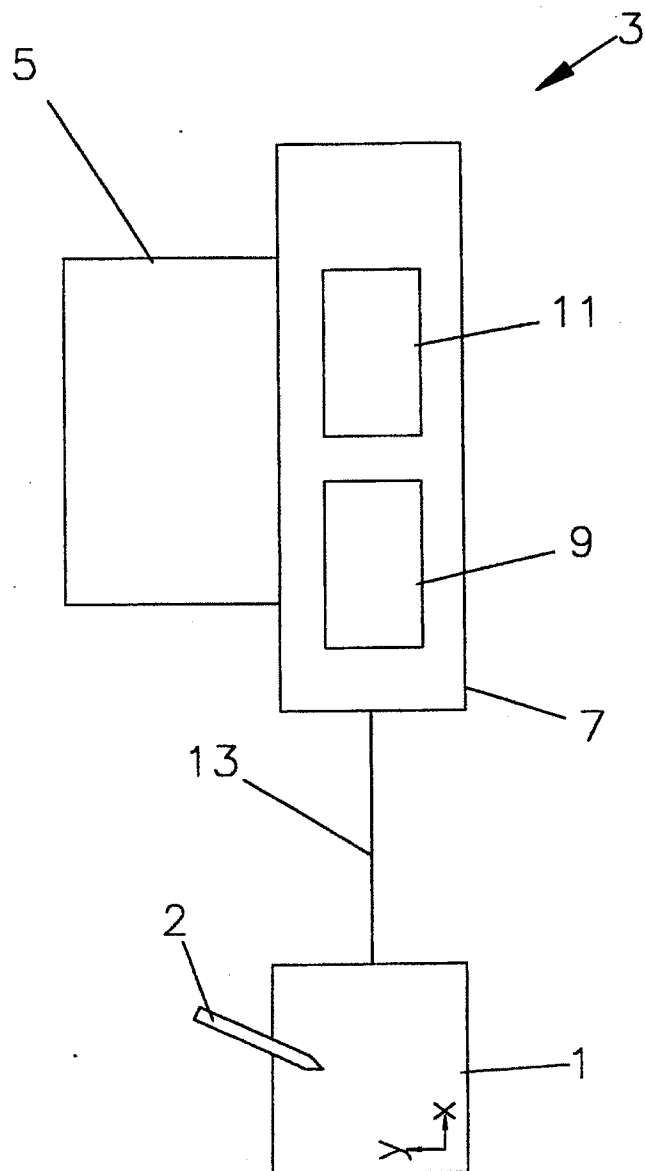
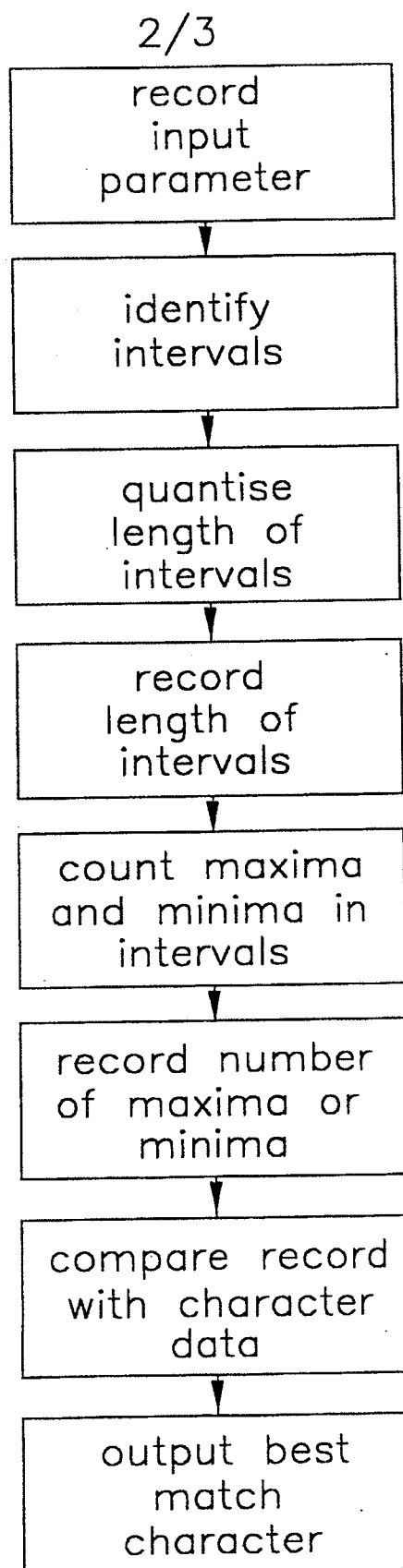


Fig 2.



3/3

Fig. 3

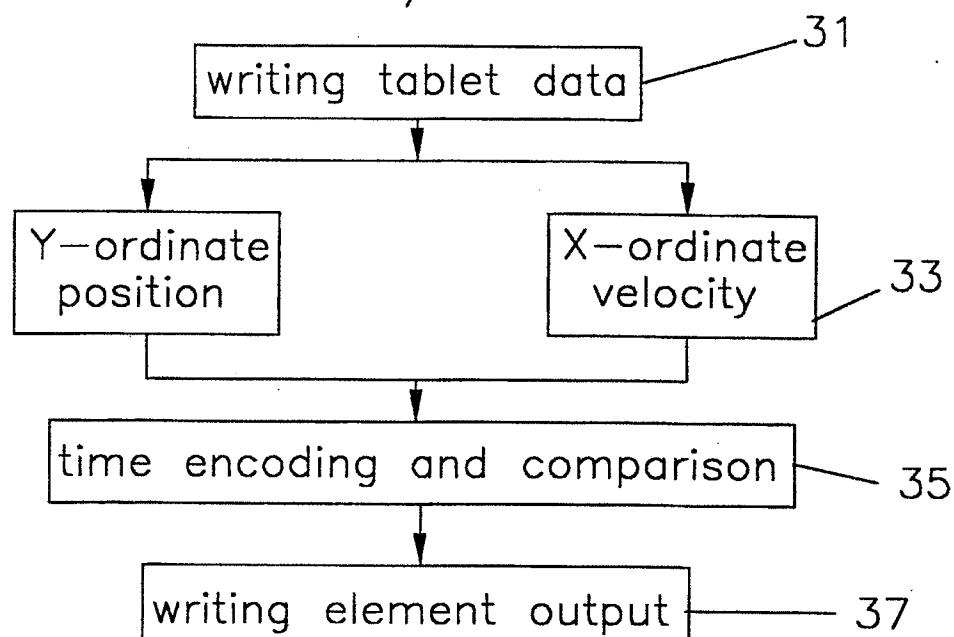
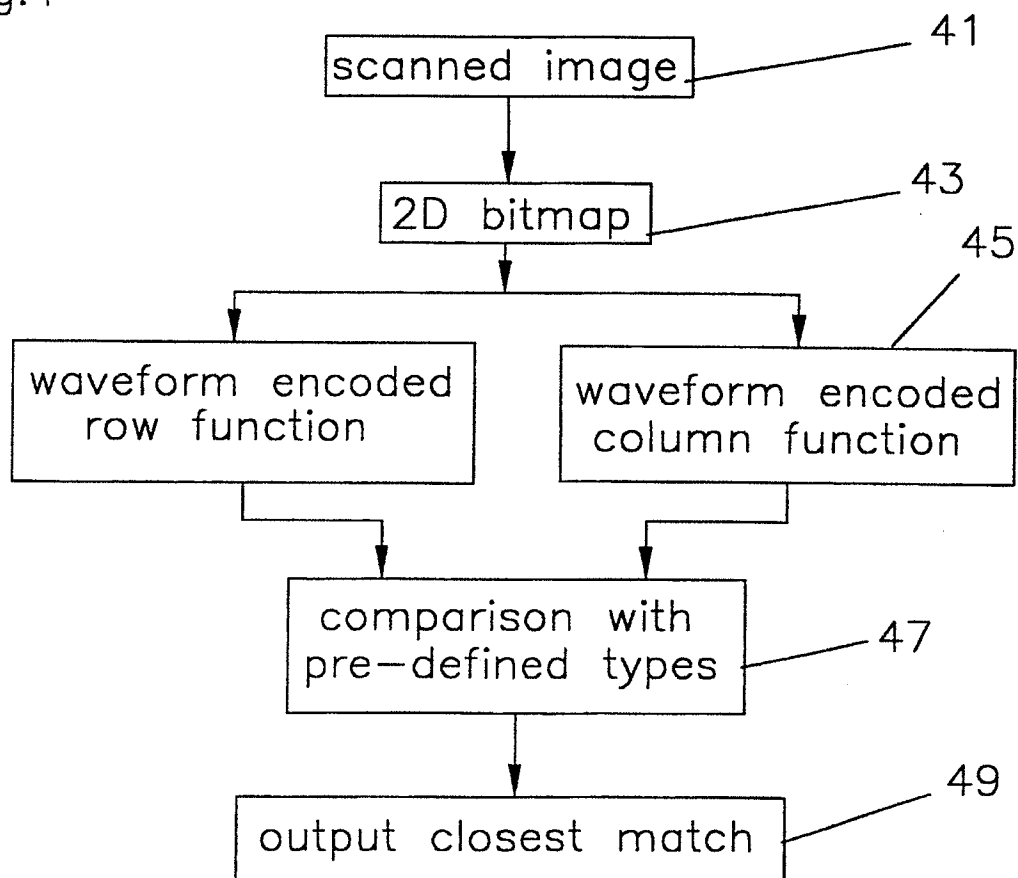


Fig.4



INTERNATIONAL SEARCH REPORT

International Application No
PCT/GB 00/01340

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 G06K9/22 G06K9/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 G06K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, INSPEC, WPI Data, PAJ, COMPENDEX

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 3 818 443 A (RADCLIFFE A) 18 June 1974 (1974-06-18) column 2, line 15 - line 54 column 6, line 58 - column 7, line 53; figures 1-7; table A ----	1-4, 8-11, 13-19
A	GORSE D ET AL: "A modular pram architecture for the classification of TESPAR-encoded speech signals" NEUROFUZZY'96. IEEE EUROPEAN WORKSHOP, PRAGUE, CZECH REPUBLIC, 16-18 APRIL 1996, vol. 6, no. 3, pages 299-304, XP000925748 Neural Network World, 1996, UIVT AV CR - NNW, Czech Republic ISSN: 1210-0552 the whole document ----- -/--	1-19

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents:

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- "&" document member of the same patent family

Date of the actual completion of the international search

28 July 2000

Date of mailing of the international search report

10/08/2000

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Granger, B

INTERNATIONAL SEARCH REPORT

International Application No
PCT/GB 00/01340

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 97 31368 A (DOMAIN DYNAMICS LTD ;KING REGINALD ALFRED (GB)) 28 August 1997 (1997-08-28) page 6, paragraph 29 -page 9, paragraph 16; figure 34 ---	1-19
A	GB 2 062 323 A (NAT RES DEV) 20 May 1981 (1981-05-20) abstract; figure 5 ---	1-19
A	PAQUET T ET AL: "RECOGNITION OF HANDWRITTEN SENTENCES USING A RESTRICTED LEXICON" PATTERN RECOGNITION,US,PERGAMON PRESS INC. ELMSFORD, N.Y, vol. 26, no. 3, 1 March 1993 (1993-03-01), pages 391-407, XP000367312 ISSN: 0031-3203 abstract -----	12

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/GB 00/01340

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
US 3818443	A	18-06-1974	NONE	
WO 9731368	A	28-08-1997	AT 188063 T	15-01-2000
			AU 1804797 A	10-09-1997
			DE 69700987 D	27-01-2000
			EP 0882288 A	09-12-1998
			JP 2000504857 T	18-04-2000
GB 2062323	A	20-05-1981	US 4369431 A	18-01-1983

Form PCT/ISA/210 (patent family annex) (July 1992)